



**HAL**  
open science

## La psychoacoustique dévoile le potentiel musical du timbre

Stephen Mcadams, Bruno L. Giordano

► **To cite this version:**

Stephen Mcadams, Bruno L. Giordano. La psychoacoustique dévoile le potentiel musical du timbre. 10ème Congrès Français d'Acoustique, Apr 2010, Lyon, France. hal-00552284

**HAL Id: hal-00552284**

**<https://hal.science/hal-00552284v1>**

Submitted on 5 Jan 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# 10ème Congrès Français d'Acoustique

Lyon, 12-16 Avril 2010

## La psychoacoustique dévoile le potentiel musical du timbre

Stephen McAdams, Bruno L. Giordano

CIRMMT, Ecole de musique Schulich, Université McGill, 555 rue Sherbrooke ouest, Montréal, Québec, Canada H3A 1E3

Le timbre est un attribut musical qui n'a pas été analysé dans la théorie musicale autant que la durée et la hauteur. Pour ces derniers, de véritables théories abondent, qui ont donné lieu aux hypothèses et à l'expérimentation sur les structures psychologiques sous-tendant le traitement des informations musicales. En revanche, l'examen des textes sur l'instrumentation et l'orchestration montre qu'ils ne sont pas basés sur une théorie du timbre. Ces traités représentent pour la plupart des recueils de recettes plutôt que des théories fondées sur des principes généraux et généralisables. L'étude de la perception du timbre a été abordée pour la première fois dans le domaine de la psychoacoustique et ces recherches commencent à s'étendre vers la perception et la cognition musicales. Les premières études psychoacoustiques du timbre ont déterminé la représentation perceptive multidimensionnelle des sons isolés (appelé « espace de timbres »), ainsi que les propriétés acoustiques qui la sous-tendent. Cette représentation soulève de nombreuses questions concernant le rôle du timbre dans la musique. Un modèle d'un espace de timbres peut-il prédire la perception de relations abstraites entre timbres en termes d'intervalles et de contours? Ces relations restent-elles perceptibles si les intervalles et contours sont « transposés » dans l'espace perceptif? L'espace de timbres est-il utile pour déterminer comment une séquence de timbres sera organisée perceptivement en flux auditifs? Les auditeurs peuvent-ils apprendre des « grammaires » de timbres par la simple exposition à un langage musical spécifique, sans apprentissage explicite? Et finalement, le timbre peut-il jouer un rôle dans des aspects de la forme musicale à grande échelle, tels que les alternances entre tension et détente musicales? Les recherches récentes montrent que les réponses à ces questions sont bien positives. La confrontation entre les résultats des expériences perceptives et cognitives, d'un côté, et les analyses musicales, de l'autre, devrait pouvoir nous conduire vers une véritable théorie musicale du timbre fondée sur la recherche en acoustique psychologique.

## 1 Introduction

« Timbre » est un mot dont l'aspect vague peut nous faire penser, à tort, que la signification est simple. Mais ce mot englobe un ensemble très complexe d'attributs auditifs, ainsi qu'une pléthore de problématiques psychologiques et musicales. Il comprend plusieurs paramètres perceptifs dont ne rendent pas compte ceux de la hauteur tonale, la sonie (ou l'intensité perçue), la position dans l'espace, la durée, ou des caractéristiques environnementales diverses comme la réverbération de la salle. Ceci laisse une richesse de possibilités que les chercheurs en psychoacoustique ont exploré depuis au moins les 4 dernières décennies. Nous savons maintenant que le timbre possède deux caractéristiques globales qui contribuent à la perception de la musique : 1) il est un ensemble très varié d'attributs sensoriels abstraits, dont certains varient au cours du temps (p.ex., le mordant de l'attaque, la brillance, la nasalité, la richesse sonore), et d'autres sont discrets ou catégoriels (p.ex., le « blatte » au début d'un *sforzando* joué par un trombone ou l'extinction brusque d'un son de clavecin), et 2) il est un des véhicules perceptifs principaux pour la reconnaissance, l'identification et le suivi au cours du temps d'une source sonore (une voix de chanteuse, une clarinette, un ensemble de cloches de carillon) [1, 2, 3]. L'approche psychoacoustique comprend également des recherches sur le timbre comme ensemble de dimensions porteuses de forme en musique [4].

## 2 Le timbre comme ensemble multidimensionnel d'attributs auditifs

Une des approches principales de la perception du timbre tente de caractériser quantitativement les différences perceptives entre sons. Des recherches précoces sur le timbre ont focalisé sur des aspects préconçus comme les poids relatifs des différentes fréquences présentes dans un son, ou sa « couleur sonore » [5]. Un voix qui chante un do tout en faisant varier la voyelle chantée, ou un joueur d'un instrument de cuivre qui tient une note en faisant varier l'embouchure et la forme de sa cavité vocale, varient tous deux la forme du spectre sonore qui représente le niveau de chaque partiel du son en fonction de sa fréquence (voir [6]). Helmholtz [7] a inventé des dispositifs résonants d'une grande ingéniosité pour contrôler la forme spectrale afin d'explorer ces aspects du timbre. Néanmoins, l'avancement marqué dans la compréhension de la représentation perceptive du timbre a dû attendre le développement des techniques d'analyse multidimensionnelle des données de proximité dans les années 60, ainsi que les techniques de traitement de signaux des années 70. Plomp [8] et Wessel [9] ont été les premiers à appliquer ces techniques à la perception du timbre.

### 2.1 L'espace de timbres

L'analyse multidimensionnelle n'impose aucune préconception sur la structure physique ou perceptive du timbre. Les auditeurs évaluent simplement toutes les paires

tirées d'un ensemble de sons sur une échelle allant de très similaire à très dissemblable. Les sons sont normalement égalisés en hauteur, sonie et durée afin que seul le timbre varie, focalisant ainsi l'attention des auditeurs sur cet ensemble de propriétés des sons. Ensuite on ajuste un modèle de distance aux évaluations de dissemblance : les sons proches dans le modèle sont similaires au niveau de leurs timbres et les timbres différents sont éloignés. On appelle la représentation graphique de cette structure un « espace de timbres ». Le modèle de base est exprimé en termes de dimensions continues partagées par l'ensemble des timbres, la présomption sous-jacente étant que tous les auditeurs utilisent les mêmes dimensions perceptives pour comparer les timbres. Des modèles plus complexes comprennent en outre des dimensions ou traits spécifiques à certains timbres (appelés « spécificités ») ainsi que des poids perceptifs différents accordés aux dimensions et spécificités par des individus ou classes d'individus [10, 11]. De telles techniques ont été appliquées aux sons synthétiques [8, 12], aux sons d'instruments resynthétisés ou simulés [10, 11, 13, 14], aux sons d'instruments enregistrés [15, 16], et même aux dyades de sons d'instruments enregistrés [17].

On trouve souvent des spécificités pour les sons acoustiques et synthétiques complexes. On considère qu'elles représentent la présence de traits uniques qui distinguent un son de tous les autres dans un contexte donné. Par exemple, dans un ensemble de sons de cuivres, de bois et de cordes frottées, un clavecin possède un trait qui n'est partagé avec aucun autre son : le retour du sautereau qui crée un bruit sourd et qui éteint rapidement le son à la fin. Ces caractéristiques pourraient apparaître comme une spécificité forte dans le modèle de distance [11, 13].

Les modèles intègrent des différences entre individus et classes comme des pondérations sur les différentes dimensions et sur l'ensemble des spécificités. Certains auditeurs portent plus d'attention aux propriétés spectrales et ignorent les aspects temporels, tandis que d'autres font l'inverse. Cette variabilité pourrait refléter soit des différences dans le traitement sensoriel, soit dans les stratégies d'écoute et d'évaluation. Il est curieux qu'aucune étude de nos jours n'ait pu étayer l'hypothèse selon laquelle les différences individuelles auraient quelque chose à voir avec l'expérience et la formation musicales [11]. Le fait que la perception du timbre soit si étroitement alliée à la capacité à reconnaître des sources sonores dans la vie quotidienne pourrait faire en sorte que tout le monde soit un expert en quelque sorte.

## 2.2 Les corrélats acoustiques des dimensions du timbre

Plusieurs études ont déterminé les corrélats acoustiques indépendants des dimensions continues du timbre en corrélant la position le long de la dimension perceptive avec un paramètre acoustique unidimensionnel calculé à partir des sons [18, 19]. Les corrélats (aussi appelés « descripteurs ») les plus omniprésents dérivés des sons d'instruments de musique comprennent le centroïde spectral (qui représente le poids relatif des fréquences aiguës et graves et qui correspond à la brillance du timbre : un hautbois a un centroïde plus élevé qu'un cor d'harmonie), le logarithme du temps d'attaque (qui distingue les sons entretenus par le souffle ou le frottement des sons d'instruments produits par le pincement ou la frappe), le

flux spectral (le degré d'évolution de la forme spectrale au cours de la durée du son, qui est élevé pour les cuivres et plus faible pour les instruments à anche simple) et l'irrégularité spectrale (l'aspect saccadé de la forme spectrale : élevé pour la clarinette et le vibraphone, faible pour la trompette). Caclin, McAdams, Smith & Winsberg [20] ont conduit une étude confirmatoire employant les jugements de dissemblance sur des sons purement synthétiques pour lesquels la nature exacte des dimensions des stimuli pouvait être contrôlée. Ils ont confirmé la perception des dimensions liées au centroïde spectral, le temps d'attaque et l'irrégularité spectrale, mais ont remis en question le statut du flux spectral.

La combinaison d'un modèle quantitatif des relations perceptives entre timbres et l'explication psychophysique des paramètres du modèle est un pas important pour atteindre un contrôle prévisible du timbre dans plusieurs domaines tels que l'analyse et la synthèse sonore et la recherche intelligente dans des bases de données sonores [21]. De telles représentations ne sont utiles que dans la mesure où elles sont : 1) généralisables au-delà de l'ensemble des sons étudiés, 2) robustes par rapport aux changements dans le contexte musical et 3) généralisables à d'autres tâches d'écoute que celles utilisées pour construire le modèle. Si une représentation possède ces propriétés, on pourrait considérer qu'elle rend compte de façon précise du timbre musical, étant caractérisé par une propriété essentielle d'un modèle scientifique, la capacité à prédire de nouveaux phénomènes empiriques.

Les modèles de l'espace de timbres ont été utiles pour prédire la perception des auditeurs dans des situations autre que celles présentées dans des expériences, suggérant ainsi que ces modèles captent des aspects importants de la représentation perceptive du timbre. Grey & Gordon [18] ont trouvé qu'en échangeant les enveloppes spectrales entre paires de sons qui se distinguaient principalement sur la dimension spectrale, chacun des deux sons a pris la place de l'autre dans l'espace de timbres. Ces espaces ont pu également prédire la perception des relations entre paires de timbres (les intervalles de timbre), ainsi que la ségrégation des flux auditifs basée sur des indices acoustiques liés au timbre (voir ci-dessous).

## 2.3 Effets du changement de hauteur sur les relations de timbre

Marozeau, de Cheveigné, McAdams & Winsberg [22] ont montré que les espaces de timbres pour les sons d'instrument de musique enregistrés sont semblables aux différentes hauteurs (si2, do#3, si-bemol3). Les auditeurs sont également capables d'ignorer les différences de hauteur au sein d'une octave lorsqu'on leur demande de comparer uniquement les timbres des sons. Si les différences de hauteur sont plus grandes d'une octave, les interactions entre les deux attributs se manifestent. Marozeau & de Cheveigné [23] ont fait varier la brillance d'un ensemble de sons de synthèse, tout en faisant varier la hauteur sur une gamme de 18 demi-tons. Ils ont trouvé que les différences de hauteur ont affecté les relations de timbre de deux façons : 1) la hauteur se trouve dans l'espace de timbres comme une dimension orthogonale aux dimensions du timbre et 2) les différences entre hauteurs affectent systématiquement la dimension du timbre lié au centroïde spectral. Ces résultats suggèrent une relation étroite entre la brillance du timbre et l'aspect grave-aigu de la hauteur. Ce

lien serait cohérent avec les représentations neurales sous-jacentes qui partagent des attributs communs, telle qu'une organisation tonotopique.

### 3 Le timbre comme véhicule pour l'identité d'une source sonore

La seconde approche du timbre concerne son rôle dans la reconnaissance de l'identité d'un instrument de musique ou plus généralement d'un événement qui produit du son. Une hypothèse raisonnable est que les dimensions sensorielles qui composent le timbre servent d'indicateurs dans la catégorisation, la reconnaissance et l'identification des événements et sources sonores [2, 24].

Les recherches sur l'identification des instruments de musique sont pertinents ici. Saldanha & Corso [25] ont étudié l'identification des sons d'instruments de musique isolés, en employant des instruments tirés de l'orchestre occidental joués avec et sans vibrato. Ils s'intéressaient à l'importance relative des transitoires d'attaque et d'extinction, de l'enveloppe spectrale de la partie entretenue du son et du vibrato. De façon étonnante, l'identification des sons isolés est très faible pour certains instruments. Lorsque les attaques et extinctions ont été coupées, l'identification a baissé de façon marquée pour certains instruments, surtout pour les attaques coupées sur des sons joués sans vibrato. Cependant, lorsque le vibrato est présent, l'effet de couper l'attaque était moindre et l'identification était meilleure. Ces résultats suggèrent que les informations importantes pour l'identification sont présentes dans l'attaque, mais qu'en l'absence de ces informations, d'autres demeurent disponibles dans la partie entretenue (bien qu'elles soient plus importantes pour certains instruments que d'autres), surtout quand le vibrato est présent. Le vibrato pourrait augmenter notre capacité à extraire des informations relatives à la structure de résonance de l'instrument [26].

Giordano & McAdams [27] ont analysé des données de la littérature sur l'identification et les jugements de dissemblance des sons musicaux. L'objectif de cette étude était de déterminer si les sons produits avec de grandes différences dans les mécanismes de production sonore créeraient des différences dans les données perceptives. Sur l'ensemble des études d'identification, les auditeurs ont confondu fréquemment les sons produits par les instruments avec des structures physiques semblables (p.ex., clarinettes et saxophones, tous deux des instruments à anche simple). Ils ont rarement confondu des sons produits par des systèmes physiques très différents (p.ex., la trompette, un instrument à anche lippale, et le basson, un instrument à anche double). La vaste majorité des espaces de timbres publiés auparavant révèlent que les sons produits avec des structures de résonance semblables (instruments à cordes comparés aux instruments à vent) ou avec des mécanismes d'excitation semblables (excitation impulsive, comme le piano, comparée à une excitation entretenue, comme la flûte), occupaient la même région de l'espace. Ces résultats suggèrent que les auditeurs peuvent identifier de façon fiable des grandes différences au niveau des mécanismes de production sonore, en se focalisant sur les attributs du timbre utilisés pour évaluer la dissemblance entre les sons musicaux.

Plusieurs recherches sur la perception des sons de tous les jours étendent le concept de timbre au-delà du contexte musical (voir [2, 24, 28] pour des revues de questions). Parmi celles-ci, les études sur les sons d'impact fournissent

des informations sur les attributs du timbre utiles pour la perception des propriétés des instruments à percussion : la géométrie des barres [29], le matériau des barres [30], le matériau des plaques [31], et la dureté des maillets [32]. Les facteurs de timbre pertinents pour les évaluations perceptives varient avec la tâche à accomplir. Les facteurs spectraux priment pour la perception de la géométrie [29]. Les facteurs spectrotemporels (p.ex., le taux de changement du centroïde spectral et du niveau sonore) dominent la perception des matériaux [30, 31] et des maillets [32].

La perception de l'identité d'un instrument, en dépit des variations de hauteur, est peut-être liée à l'invariance du timbre, les aspects du timbre qui restent constants avec un changement de hauteur et de sonie. Handel & Erickson [33] ont trouvé que les auditeurs sans formation musicale peuvent reconnaître deux sons produits à des hauteurs différentes comme provenant du même instrument ou de la même voix seulement dans une gamme de hauteur de l'ordre de l'octave. Steele & Williams [34] ont montré que les auditeurs avec une formation musicale peuvent exécuter cette tâche avec un taux de réussite de 80 % même avec une différence de hauteur de l'ordre de 2,5 octaves. Ces résultats suggèrent qu'il y a des limites à l'invariance du timbre à travers les hauteurs, mais que ces limites dépendent de la formation musicale.

Son rôle dans l'identification et la catégorisation des sources est peut-être l'aspect le plus négligé du timbre et apporte avec lui des avantages et désavantages pour l'emploi du timbre comme dimension porteuse de forme en musique [4]. Un des avantages est que la catégorisation et l'identification d'une source sonore pourrait mettre en jeu des connaissances perceptives (acquises par l'auditeur implicitement par son expérience dans le monde quotidien et dans des situations musicales). Ces connaissances l'aident à suivre une voix ou un instrument donné dans une texture musicale complexe. Les auditeurs font cela avec aisance et certaines recherches ont montré que les facteurs de timbre peuvent faire une contribution importante dans le suivi des voix [35, 36], une capacité qui est particulièrement utile face à des textures polyphoniques.

Les désavantages peuvent se manifester dans des situations où le compositeur cherche à créer des mélodies qui se baladent d'un instrument à l'autre, p.ex., dans les *Klangfarbenmelodien* de Schoenberg [37]. Notre prédisposition à identifier la source sonore et de la suivre au cours du temps empêcherait une perception plus relative pour laquelle les différences de timbre seraient perçues comme un mouvement à travers l'espace de timbres plutôt que comme un simple changement de source sonore. Pour des cas où de telles compositions de timbre marchent, les compositeurs ont souvent pris des précautions pour créer une situation musicale qui conduit l'auditeur vers un mode d'écoute où la perception des relations prime sur celle des objets.

### 4 Intervalles de timbre

Si on peut démontrer la possibilité de percevoir des intervalles de timbre, la porte sera ouverte pour le développement d'opérations musicales sur des séquences de timbre, de la même manière que les compositeurs font des opérations sur les intervalles de hauteurs [5]. Un autre intérêt de cette exploration est qu'il étend l'utilisation de l'espace de timbres comme un modèle au-delà du paradigme de dissemblance.

Ehresman & Wessel [14, 38] ont été les premiers à proposer cette piste, en développant une tâche dans laquelle les auditeurs devaient juger la similarité d'intervalles formés par des paires de timbres. L'idée de base est que les intervalles de timbres auraient des propriétés perceptives semblables à celles des intervalles de hauteurs : un intervalle de hauteurs est une relation le long d'une dimension bien ordonnée qui retient une certaine invariance sous certains types de transformation, tels que la translation sur une dimension, ce que les musiciens appellent la « transposition ». Mais quelle pourrait être la signification d'une transposition dans un espace multidimensionnel ? On peut considérer un intervalle de timbre comme un vecteur dans l'espace qui lie deux timbres. Il a une longueur précise (la distance entre les timbres) et une orientation précise. Prises ensemble, ces deux propriétés définissent la quantité de changement le long de chaque dimension de l'espace nécessaire pour passer d'un timbre à un autre. Si on présume que ces dimensions sont continues et linéaires du point de vue perceptif, les paires de timbres caractérisées par la même relation vectorielle devraient avoir la même relation perceptive et former ainsi le même intervalle de timbres. La transposition dans ce cas consiste à translater le vecteur n'importe où ailleurs dans l'espace, tant que sa longueur et son orientation soient préservées.

Ehresman et Wessel ont testé cette hypothèse en utilisant une tâche dans laquelle les auditeurs devaient comparer deux intervalles de timbres (p.ex., A-B contre C-D) et ordonner plusieurs timbres D selon leur capacité à maintenir l'analogie : timbre A est à timbre B ce que timbre C est à timbre D. Ils ont trouvé que plus le timbre D était proche d'un point idéal défini par le modèle vectoriel dans l'espace de timbres (c'est-à-dire que le vecteur C-D idéal était une simple translation du vecteur A-B), plus le vecteur C-D était préféré.

Par la suite, McAdams & Cunibile [39] ont testé le modèle vectoriel dans l'espace 3D de Krumhansl [13] (en ignorant les spécificités). Cinq ensembles de timbres à des endroits différents de l'espace de timbres ont été choisis pour chaque comparaison afin de voir si les résultats se généralisaient. Les non musiciens et les compositeurs de musique électroacoustique participaient à l'étude pour voir si la formation et l'expérience musicales avait un effet. Les auditeurs ont trouvé la tâche assez difficile, ce qui n'est pas très surprenant étant donné que même les compositeurs professionnels n'ont pratiquement aucune expérience avec des musiques qui utilisent les intervalles de timbres de façon systématique. Le résultat principal est encourageant dans le sens où les données étaient globalement le modèle vectoriel, bien que celles des compositeurs conformaient plus au modèle que celles des non musiciens. Cependant, lorsque l'on examine de près les 5 versions différentes pour chaque type de comparaison, il est clair que certaines comparaisons de timbres ne vont pas dans le sens des prédictions du modèle.

Un facteur à prendre en considération est que les spécificités sur certains timbres dans cet ensemble ont été ignorées. Et ces spécificités distordraient nécessairement les vecteurs utilisés pour choisir les timbres, car les spécificités sont comme une dimension supplémentaire pour chaque timbre. Ainsi, certains intervalles de timbres correspondent bien à ce qui est prédit parce que les spécificités sont absentes ou de valeur faible, tandis d'autres seraient sérieusement distordus et donc pas perçus comme semblables aux autres intervalles du fait de spécificités

élevées. Ce genre de raisonnement suggère que l'utilisation des intervalles de timbres comme partie intégrante d'un discours musical court le risque d'être difficile à accomplir avec des sources sonores très complexes et idiosyncrasiques, parce qu'ils auront vraisemblablement des spécificités. L'utilisation des intervalles de timbres pourrait, à long terme, être limitée aux sons synthétisés ou des sons mélangés créés par la combinaison de plusieurs instruments.

## 5 Le timbre et le groupement musical

Une façon importante dont le timbre peut contribuer à l'organisation d'une structure musicale est liée au fait que les auditeurs tendent à connecter perceptivement les événements successifs provenant de la même source sonore. En général, une source donnée produira des sons qui sont relativement semblables en termes de hauteur, sonie, timbre et position spatiale d'un événement à l'autre (voir [40, 41] pour des revues de questions). On appelle la liaison perceptive d'événements sonores successifs en un « message » cohérent au cours du temps « l'intégration en flux auditifs », et la séparation des événements en messages distincts « la ségrégation des flux auditifs » [42]. Un des principes de base qui semble opérer dans la formation de flux auditifs est le suivant : des événements successifs qui sont relativement semblables au niveau de leurs propriétés spectrotemporelles (c'est-à-dire de leurs timbres et hauteurs) pourraient provenir de la même source sonore et devraient être groupés ensemble ; des sources individuelles ne tendent pas à changer les propriétés acoustiques soudainement et de façon répétée d'un événement à l'autre. Les premières démonstrations de la formation des flux sur la base du timbre [14] suggèrent un lien entre la représentation de l'espace de timbres et la tendance à former des flux sur la base des différences spectrales résultant des différences de timbre [41]. Les premiers chercheurs ont été convaincus que c'était principalement les aspects spectraux du timbre (comme le centroïde spectral) qui étaient responsables de la ségrégation des flux et que les aspects temporels (comme le temps d'attaque) avaient peu d'effet [43]

Récemment les opinions ont changé significativement et plusieurs études indiquent un rôle important pour les attributs à la fois spectraux et temporels du timbre dans la ségrégation auditive [44]. Iverson [45] a utilisé des séquences alternant entre deux sons d'instruments enregistrés ayant les mêmes hauteurs et sonies et a demandé aux auditeurs d'évaluer le degré de ségrégation perçue. Une analyse multidimensionnelle des jugements de ségrégation (traités comme une mesure de dissemblance) a été effectuée pour déterminer quels attributs acoustiques contribuaient à l'impression de ségrégation auditive. Une comparaison avec un espace de timbres construits avec les mêmes sons [15] a montré que les attributs acoustiques à la fois spectraux (centroïde spectral) et dynamiques (temps d'attaque et flux spectral) étaient impliqués dans la ségrégation. D'autres résultats étaient cette hypothèse [46, 47].

Tous ces résultats sont importants pour la théorie de l'analyse des scènes auditives. D'un côté, ils montrent que plusieurs des propriétés acoustiques d'une source sont prises en compte dans la formation des flux auditifs. D'un autre côté, ils sont importants pour faire de la musique (que ce soit avec des synthétiseurs ou des instruments acoustiques), car ils montrent que plusieurs aspects du timbre affectent fortement l'organisation de la surface musicale en flux

auditifs. Des orchestrations différentes d'une séquence de hauteurs donnée peut changer complètement ce qui est entendu comme mélodie et rythme, comme l'a montré Wessel [14]. Le timbre est également un composant important dans la perception des groupements musicaux, qu'ils soient au niveau des séquences de notes distinguées par des changements de timbre [48] et des plus grandes sections musicales délimitées par des changements marqués d'orchestration ou de texture timbrale [49].

## 6 Le timbre comme force structurante dans la musique

La perception du timbre est au cœur de l'orchestration, un domaine de la pratique musicale qui n'a guère été étudiée expérimentalement. Des combinaisons d'instruments peuvent donner lieu à de nouveaux timbres si les sons sont perçus comme mélangés ou fusionnés perceptivement. Et le timbre peut jouer un rôle dans la création et le relâchement de tension musicale.

### 6.1 Mélanges de timbres

La création de nouveaux timbres par l'orchestration dépend nécessairement du degré auquel les sources sonores constituantes se fusionnent ou se mélangent pour créer le nouveau son émergent [50, 51]. Sandell [52] a proposé trois classes d'objectifs perceptifs dans la combinaison des instruments : l'*hétérogénéité timbrale* dans laquelle on cherche à maintenir les instruments perceptivement distincts, l'*augmentation timbrale* dans laquelle un instrument embellit un autre qui domine perceptivement la combinaison, et l'*émergence timbrale* dans laquelle un nouveau son en résulte qui n'est identifié comme aucun des deux constituants. Le mélange semble dépendre d'un certain nombre de facteurs acoustiques tels que le synchronisme des attaques des sons et d'autres qui sont plus directement liés au timbre, comme la similarité des temps d'attaque et des centroïdes spectraux, ainsi que le centroïde global de la combinaison. Lorsque l'on combine des sons entretenus et percussifs, le son percussif a l'effet la plus forte sur leur fusion perceptive, mais les propriétés spectrales du son entretenu domine la perception du timbre du mélange [53].

### 6.2 L'apprentissage implicite des grammaires de timbre

Pour utiliser le timbre de manière syntaxique dans la musique, il faudrait que les auditeurs puissent apprendre des règles d'enchaînement de timbres comme pour le durée et la hauteur. Cette possibilité a été exploré par Bigand, Perruchet & Boyer [54] dans un premier temps, suivi par les études de Tillmann & McAdams [55] par la suite.

Bigand et al. [54] ont présenté des grammaires artificielles de sons musicaux. Des règles d'enchaînement ont été créées. Après une exposition aux séquences construites avec la grammaire, les auditeurs entendaient de nouvelles séquences et devaient dire si chaque séquence était conforme ou non à la grammaire apprise. Le taux de bonnes réponses était supérieur au hasard, indiquant la capacité des auditeurs à apprendre une grammaire de timbre.

Tillmann & McAdams [55] ont étudié l'influence des caractéristiques acoustiques sur l'apprentissage implicite de régularités statistiques (les probabilités de transition) dans des séquences de timbres musicaux. Ces régularités formaient des groupes de trois timbres successifs (tiré de

[11]) : la probabilité de transition entre le premier et deuxième ou entre le deuxième et troisième timbres était largement supérieure à celle entre le troisième timbre d'un groupe et le premier timbre de tous les autres groupes dans l'expérience. Dans la phase d'apprentissage implicite, les auditeurs entendaient une séquence isochrone de timbres pendant une vingtaine de minutes. En outre, les séquences ont été construites pour que le groupement auditif sur la base de similarités entre les timbres soit conforme ou pas à la structure suggérée par les régularités statistiques. Ensuite des groupes de trois timbres qui faisaient parti de la grammaire ou qui n'en faisaient pas ont été présentés et l'auditeur devait les classer comme conforme ou non aux régularités. Les auditeurs étaient capables d'apprendre la grammaire, mais curieusement cet apprentissage ne dépendait pas de la conformité entre la structure du groupement sur la base des caractéristiques acoustiques et les régularités statistique réglant les probabilités de transition entre timbres.

### 6.3 Le rôle du timbre dans la tension musicale

Le timbre peut également contribuer à la forme musicale à plus grande échelle et, plus spécifiquement, au sens du mouvement entre la tension et la détente. Ce mouvement a été considéré par bon nombre de théoriciens de la musique comme une des bases primaires de la perception de la grande forme en musique. Il a été associé traditionnellement à l'harmonie dans la musique occidentale et joue un rôle important dans la théorie générative de la musique tonale de Lerdahl & Jackendoff [56]. Les travaux expérimentaux sur le rôle de l'harmonie dans la perception de la tension et la détente musicales (ou inversement, la sensation de tension qui accompagne un moment auquel la musique doit continuer et la sensation de détente qui accompagne l'achèvement d'une phrase musicale) a suggéré que la rugosité auditive est un composant important de la tension perçue [57]. La rugosité est un attribut élémentaire du timbre fondé sur la sensation de fluctuations rapides dans l'enveloppe d'amplitude. Elle peut être générée par des composantes fréquentielles proximales dont l'interaction dans la périphérie auditive crée des battements. Des intervalles dissonants tendent à avoir plus de tels battements que les intervalles consonants. Ainsi, une relation assez directe entre la dissonance sensorielle et la rugosité a été mise en évidence (voir [58, 59] pour des revues de questions).

Comme première étape vers une compréhension de comment ceci pourrait opérer dans la musique, Paraskeva & McAdams [60] ont mesuré l'inflexion de la tension et la détente musicales résultant d'un changement de timbre. Les auditeurs devaient évaluer sur une échelle en 7 points le degré d'achèvement perçu de la musique à plusieurs endroits où la musique a été arrêtée. Ce qui en résulte est un profil d'achèvement, qui peut être utilisé pour inférer la tension musicale en établissant une équivalence entre l'achèvement et la détente, d'un côté, et le manque d'achèvement et la tension, de l'autre. Nous avons testé deux morceaux de musique : un extrait du *Ricercar* tiré de l'*Offrande Musicale* pour 6 voix de Bach (œuvre tonale) et le premier mouvement des *Six Pièces pour Orchestre* de Webern (œuvre non tonale). Chaque extrait a été joué par un échantillonneur numérique en version orchestrale (l'orchestration par Webern de l'œuvre de Bach a été utilisée) et en version pour piano (une transcription directe

des versions orchestrales). Il y avait des différences significatives entre les deux versions, témoignant ainsi d'un effet notoire du timbre sur la tension musicale perçue. Cependant, aux moments où les versions étaient différentes dans les œuvres, la version orchestrale était toujours plus détendue que la version pour piano.

L'hypothèse posée par Paraskeva & McAdams [60] pour cet effet était que la plus grande détente de la version orchestrale pourrait être due aux processus impliqués dans la formation des flux auditifs et à la dépendance de la rugosité perçue sur les résultats de tels processus d'organisation perceptive [61]. La rugosité, ou tout autre attribut d'un événement auditif isolé, est calculé après que les processus d'organisation auditive ont groupé ensemble les informations auditives. Les sons de piano ont une attaque assez franche. Si plusieurs notes ont lieu en même temps dans la partition et sont jouées avec un son de piano, elles seront très synchrones. Puisqu'elles commencent en même temps et ont des enveloppes d'amplitude semblables, elles tendraient à se fusionner perceptivement et la rugosité calculée résulterait de l'interaction de toutes les composantes fréquentielles de toutes les notes.

La situation en est autre pour la version orchestrale pour deux raisons. La première est que pour des raisons de contrôle expérimental, le même timing a été utilisé pour les versions pour piano et pour orchestre. Dans ce dernier, plusieurs instruments qui jouent ont des attaques lentes, tandis que d'autres ont des attaques rapides. Il pourrait y avoir donc une grande asynchronisme entre les instruments en termes du temps d'attaque perçu [62]. En outre, puisque les timbres de ces instruments sont souvent très différents, plusieurs voix différentes avec des timbres différents arrivent de façon momentanée à une sonorité verticale donnée, mais la verticalité n'est pas perçue car l'auditeur continuerait vraisemblablement à suivre perceptivement les instruments individuels horizontalement, privilégiant ainsi la ségrégation des flux plutôt que leur intégration. Ainsi, l'asynchronisme des attaques et la décomposition des verticalités en horizontalités concourraient à réduire le degré de fusion perceptive. Une fusion réduite signifie une plus grande ségrégation. Et alors la rugosité de la version orchestrale serait calculée sur chaque événement auditif seul plutôt que sur la masse sonore composée de l'ensemble des événements simultanés. Ces rugosités individuelles dans la version orchestrale seraient vraisemblablement moindres que celles de la version pour piano. Donc une fois de plus, la composition timbrale peut avoir une interaction étroite avec les processus de formation des flux auditifs.

## 7 Conclusion

Le timbre musical est une combinaison de dimensions perceptives continues et de traits discrets auxquels les auditeurs sont sensibles mais avec des degrés de sensibilité différents selon l'individu et la dimension. Les dimensions continues ont souvent des corrélats acoustiques quantifiables. La représentation de l'espace de timbres est un modèle psychologique puissant qui permet de prédire des aspects de la perception du timbre dans des situations au-delà de celles utilisées pour dériver le modèle en première instance. Les intervalles de timbre, par exemple, peuvent être conçus comme des vecteurs au sein de l'espace. L'espace de timbres peut également prédire qualitativement la grandeur des différences de timbre qui provoquerait la ségrégation des flux auditifs.

Le choix des timbres pour former des mélanges affecte le degré de fusion perceptive et la perception du timbre émergeant d'un mélange. Les auditeurs sont capables d'apprendre implicitement, c'est-à-dire sans formation explicite, les grammaires construites sur les règles d'enchaînement des timbres. Le timbre peut jouer un rôle dans des mouvements de tension et de détente musicales et ainsi contribuer à l'expression inhérente de la forme musicale. Dans des conditions de fort mélange entre des instruments qui composent une sonorité verticale, la rugosité timbrale est un composant majeur de la tension musicale. Cependant, elle dépend fortement de la façon dont les processus de groupement auditif ont analysé les informations acoustiques entrant en événements et en flux. Ainsi, la psychoacoustique et la psychologie cognitive apportent bon nombre de possibilités pour l'utilisation du timbre dans la musique. Il reste maintenant aux compositeurs d'en profiter dans la composition de leurs œuvres !

## Remerciements

Les recherches résumées dans cet article ont été menées à l'Ircam-Centre Pompidou (SMTS-CNRS), au Laboratoire de Psychologie Expérimentale (CNRS-Université Paris V) et à l'Université McGill. Les recherches récentes ont bénéficié d'un soutien du Conseil de recherches en sciences naturelles et en génie du Canada et du programme des Chaires de recherche du Canada. Cet article est une traduction modifiée et étendue de [63], qui est lui-même une mise à jour de [64]. Je remercie Michel Vallières pour son aide précieuse dans la relecture du texte français.

## Références

- [1] Hajda, J. M., Kendall, R. A., Carterette, E. C. & Harshberger, M. L. (1997). Methodological issues in timbre research. In I. Deliège & J. Sloboda (Eds.), *Perception and Cognition of Music* (pp. 253-306). Hove: Psychology Press.
- [2] McAdams, S. (1993). Recognition of sound sources and events. In S. McAdams & E. Bigand (Eds.), *Thinking in Sound: The Cognitive Psychology of Human Audition* (pp. 146-198). Oxford: Oxford University Press.
- [3] Risset, J.-C. & Wessel, D. L. (1999). Exploration of timbre by analysis and synthesis. In D. Deutsch (Ed.), *The Psychology of Music* (2nd ed., pp. 113-168). San Diego: Academic Press.
- [4] McAdams, S. (1989). Psychological constraints on form-bearing dimensions in music. *Contemporary Music Review*, 4(1), 181-198.
- [5] Slawson, W. (1985). *Sound Color*. Berkeley, CA: University of California Press.
- [6] McAdams, S., Depalle, P. & Clarke, E. (2004b). Analyzing musical sound. In E. Clarke & N. Cook (Eds.), *Empirical Musicology: Aims, Methods, Prospects* (pp. 157-196). New York: Oxford University Press.
- [7] Helmholtz, H. L. F. von (1885/1954). *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. New York, from 1877 trans. by A. J. Ellis of 4th German ed.; republ. 1954 by Dover.

- [8] Plomp, R. (1970). Timbre as a multidimensional attribute of complex tones. In R. Plomp & G. F. Smoorenburg (Eds.), *Frequency Analysis and Periodicity Detection in Hearing* (pp. 397-414). Leiden: Sijthoff.
- [9] Wessel, D. L. (1973). Psychoacoustics and music: A report from Michigan State University. *PACE: Bulletin of the Computer Arts Society*, **30**, 1-2.
- [10] Grey, J. M. (1977). Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America*, **61**, 1270-1277.
- [11] McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G. & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. *Psychological Research*, **58**, 177-192.
- [12] Miller, J. R. & Carterette, E. C. (1975). Perceptual space for musical structures. *Journal of the Acoustical Society of America*, **58**, 711-720.
- [13] Krumhansl, C. L. (1989). Why is musical timbre so hard to understand? In S. Nielzén & O. Olsson (Eds.), *Structure and Perception of Electroacoustic Sound and Music* (pp. 43-53). Amsterdam: Excerpta Medica.
- [14] Wessel, D. L. (1979). Timbre space as a musical control structure. *Computer Music Journal*, **3**(2), 45-52.
- [15] Iverson, P. & Krumhansl, C. L. (1993). Isolating the dynamic attributes of musical timbre. *Journal of the Acoustical Society of America*, **94**, 2595-2603.
- [16] Lakatos, S. (2000). A common perceptual space for harmonic and percussive timbres. *Perception & Psychophysics*, **62**, 1426-1439.
- [17] Kendall, R. A. & Carterette, E. C. (1991). Perceptual scaling of simultaneous wind instrument timbres. *Music Perception*, **8**, 369-404.
- [18] Grey, J. M. & Gordon, J. W. (1978). Perceptual effects of spectral modifications on musical timbres. *Journal of the Acoustical Society of America*, **63**, 1493-1500.
- [19] Krimphoff, J., McAdams, S. & Winsberg, S. (1994). Caractérisation du timbre des sons complexes. II: Analyses acoustiques et quantification psychophysique [Characterization of the timbre of complex sounds. II: Acoustic analyses and psychophysical quantification]. *Journal de Physique*, **4**(C5), 625-628.
- [20] Caclin, A., McAdams, S., Smith, B. K. & Winsberg, S. (2005). Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones. *Journal of the Acoustical Society of America*, **118**, 471-482.
- [21] Peeters, G., McAdams, S. & Herrera, P. (2000) Instrument sound description in the context of MPEG-7. *Proceedings of the 2000 International Computer Music Conference, Berlin* (pp. 166-169). San Francisco: International Computer Music Association.
- [22] Marozeau, J., de Cheveigné, A., McAdams, S. & Winsberg, S. (2003). The dependency of timbre on fundamental frequency. *Journal of the Acoustical Society of America*, **114**, 2946-2957.
- [23] Marozeau, J. & de Cheveigné, A. (2007). The effect of fundamental frequency on the brightness dimension of timbre. *Journal of the Acoustical Society of America*, **121**, 383-387.
- [24] Handel, S. (1995). Timbre perception and auditory object identification. In B. Moore (Ed.), *Hearing* (pp. 425-462). San Diego, CA: Academic Press.
- [25] Saldanha, E. L. & Corso, J.F. 1964. Timbre cues and the identification of musical instruments. *Journal of the Acoustical Society of America*, **36**, 2021-2126.
- [26] McAdams, S. & Rodet, X. (1988) The role of FM-induced AM in dynamic spectral profile analysis. In H. Duifhuis, J.W. Horst & H.P. Wit (Eds.) *Basic Issues in Hearing*, Academic Press, Londres, pp. 359-369.
- [27] Giordano, B. L. & McAdams, S. (in press). Sound source mechanics and musical timbre perception: Evidence from previous studies. *Music Perception*.
- [28] Lutfi, R. (2008). Human sound source identification. In W. Yost, A. Popper & R. Fay (Eds.), *Auditory Perception of Sound Sources* (pp. 13-42). New York: Springer.
- [29] Lakatos, S., McAdams, S. & Caussé, R. (1997). The representation of auditory source characteristics: simple geometric form. *Perception & Psychophysics*, **59**, 1180-1190.
- [30] McAdams, S., Chaigne, A. & Roussarie, V. (2004a). The psychomechanics of simulated sound sources: Material properties of impacted bars. *Journal of the Acoustical Society of America*, **115**, 1306-1320.
- [31] Giordano, B. L. & McAdams, S. (2006). Material identification of real impact sounds: Effects of size variation in steel, glass, wood and plexiglass plates. *Journal of the Acoustical Society of America*, **119**, 1171-1181.
- [32] Freed, D. J. (1990). Auditory correlates of perceived mallet hardness for a set of recorded percussive events. *Journal of the Acoustical Society of America*, **87**, 1236-1249.
- [33] Handel, S. & Erickson, M. (2001). A rule of thumb: The bandwidth for timbre invariance is one octave. *Music Perception*, **19**, 121-126.
- [34] Steele, K. & Williams, A. (2006). Is the bandwidth for timbre invariance only one octave? *Music Perception*, **23**, 215-220.
- [35] Culling, J.F. & Darwin, C.J. 1993. The role of timbre in the segregation of simultaneous voices with intersecting Fo contours. *Perception & Psychophysics*, **34**, 303-309.
- [36] Gregory, A.H. 1994. Timbre and auditory streaming. *Music Perception*, **12**, 161-174.
- [37] Schoenberg, A. (1911/1978). *Theory of Harmony*. Original German publication, 1911 ed. University of California Press, Berkeley, CA.
- [38] Ehresman, D. & Wessel, D. L. (1978). Perception of timbral analogies. *Rapports de l'IRCAM*, vol. 13. Paris: IRCAM.
- [39] McAdams, S. & Cunibile, J. C. (1992). Perception of timbral analogies. *Philosophical Transactions of the Royal Society, London*, series



- B, **336**, 383-389.
- [40] Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press.
- [41] McAdams, S. & Bregman, A. S. (1979). Hearing musical streams. *Computer Music Journal*, **3**(4), 26-43.
- [42] Bregman, A. S. & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, **89**, 244-249.
- [43] Hartmann, W. M. & Johnson, D. (1991). Stream segregation and peripheral channeling. *Music Perception*, **9**, 155-184.
- [44] Moore, B. C. J. & Gockel, H. (2002). Factors influencing sequential stream segregation. *Acustica united with Acta Acustica*, **88**, 320-332.
- [45] Iverson, P. (1995). Auditory stream segregation by musical timbre: Effects of static and dynamic acoustic attributes. *Journal of Experimental Psychology: Human Perception and Performance*, **21**, 751-763.
- [46] Bey, C. & McAdams, S. (2003). Post-recognition of interleaved melodies as an indirect measure of auditory stream formation. *Journal of Experimental Psychology: Human Perception and Performance*, **29**, 267-279.
- [47] Singh, P. G. & Bregman, A. S. (1997). The influence of different timbre attributes on the perceptual segregation of complex-tone sequences. *Journal of the Acoustical Society of America*, **120**, 1943-1952.
- [48] Deliège, I. 1987. Grouping conditions in listening to music: An approach to Lerdahl & Jackendoff's grouping preference rules. *Music Perception*, **4**, 325-360.
- [49] Deliège, I. 1989. A perceptual approach to contemporary musical forms. *Contemporary Music Review*, **4**, 213-230.
- [50] Brant, H. (1971). Orchestration. In J. Vinton (Ed.), *Dictionary of Contemporary Music* (pp. 538-546). New York: E. P. Dutton.
- [51] Erickson, R. (1975). *Sound Structure in Music*. Berkeley, CA: University of California Press.
- [52] Sandell, G. J. (1995). Roles for spectral centroid and other factors in determining "blended" instrument pairings in orchestration. *Music Perception*, **13**, 209-246.
- [53] Tardieu, D. & McAdams, S. (en préparation). Perception of dyads of percussive and sustained instruments.
- [54] Bigand, E., Perruchet, P. & Boyer, M. (1998). Implicit learning of an artificial grammar of musical timbres. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, **17**(3), 577-600.
- [55] Tillmann, B. & McAdams, S. (2004). Implicit learning of musical timbre sequences: Statistical regularities confronted with acoustical (dis)similarities. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **30**(5), 1131-1142.
- [56] Lerdahl, F. & Jackendoff, R. (1983). *The Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.
- [57] Bigand, E., Parncutt, R., & Lerdahl, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception & Psychophysics*, **58**, 125-141.
- [58] Parncutt, R. (1989). *Harmony: A Psychoacoustical Approach*. Berlin: Springer-Verlag.
- [59] Plomp, R. (1976). *Aspects of Tone Sensation: A Psychophysical Study*. London: Academic Press.
- [60] Paraskeva, S. & McAdams, S. (1997). Influence of timbre, presence/absence of tonal hierarchy and musical training on the perception of tension/relaxation schemas of musical phrases. *Proceedings of the 1997 International Computer Music Conference, Thessaloniki* (pp. 438-441). San Francisco: International Computer Music Association.
- [61] Wright, J. K. & Bregman, A. S. (1987). Auditory stream segregation and the control of dissonance in polyphonic music. *Contemporary Music Review*, **2**(1), 63-92.
- [62] Gordon, J. W. (1987). The perceptual attack time of musical tones. *Journal of the Acoustical Society of America*, **82**, 88-105.
- [63] McAdams, S. & Giordano, B. L. (2009). The perception of musical timbre. In S. Hallam, I. Cross & M. Thaut (Eds.), *Oxford Handbook of Music Psychology*, Oxford University Press, New York, pp. 72-80.
- [64] McAdams, S. (1999). Perspectives on the contribution of timbre to musical structure. *Computer Music Journal*, **23**(2), 96-113.